DOCUMENT RESUME

ED 102 676 EA 006 797

AUTHOR	Haggart, Sue A.
TITLE	The Contribution of Demonstration Programs to Educational Policymaking.
INSTITUTION	Rand Corp., Santa Monica, Calif.
REPORT NO	P-5333
PUB DATE	Dec 74
NOTE	21p.
AVAILABLE FROM	Rand Corporation, 1700 Main Street, Santa Monica, California 90406 (Order No. P-5333, $1.50)
EDRS PRICE	MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS	*Comparative Analysis; *Cost Effectiveness; Demonstration Programs; *Demonstrations (Educational); Elementary Secondary Education; *Policy Formation; Program Costs; *Program Evaluation

ABSTRACT

Policy, or program, analysis has the responsibility for providing policymakers with timely information about the consequences of alternative courses of action. In elementary and secondary education, however, the policymaking effort is complicated by the organizational gap often present between the policymaker and the practitioner or user. An educational demonstration program can help bridge this gap if analysis of the demonstration program serves as the communication link. This paper delineates an equitable way to develop cost measures, to assess achievement measures (as one outcome), and to compare alternatives using these measures. Brief examples illustrate a simple, analytically-sound way to make the most of cost and outcome data in comparing alternative programs. Because the analysis is straightforward, the results are likely to be more credible to the policymaker. (Author)

THE CONTRIBUTION OF DEMONSTRATION PROGRAMS TO EDUCATIONAL POLICYMAKING

Sue A. Haggart

December 1974

P-5333

The Rand Paper Series

The Rand Corporation
Santa Monica, California 90406

ABSTRACT

THE CONTRIBUTION OF DEMONSTRATION PROGRAMS TO EDUCATIONAL POLICYMAKING

Policy analysis or program analysis has the responsibility for providing policymakers with timely information about the consequences of alternative courses of action. In elementary and secondary education, however, the policymaking effort is complicated by the organizational gap often present between the policymaker and the practitioner or user.

The educational demonstration program can help bridge this gap, if analysis of the demonstration program serves as the communication link. This discussion delineates an equitable way to develop cost measures, to assess achievement measures (as one outcome) and to compare alternatives using these measures.

Brief examples illustrate a simple, analytically-sound way to make the most of cost and outcome data in comparing alternative programs. Because the analysis is straightforward, the results are likely to be more credible to the policymaker.

The quality of policy decisions is related, in part, to the quality and timeliness of information about the alternative courses of action available. Policy analysis or program analysis--the assessment of the means to effect the policy--has the responsibility for providing the required information. In many areas of elementary and secondary education, the policymaking effort is complicated by the organizational gap often present between policymakers and those in the action arena. Policymakers, who are remote from field operations, need a communication link with the practitioners or users. This is particularly true when assessments of, and commitments to, innovative instructional activities are to be made--activities which are frequently tested in demonstration programs.

This discussion proposes that educational demonstration programs can help bridge this gap if analysis of the demonstration programs serves as the communication link. It is hoped that the discussion will show how generally-accepted, analytically-powerful, cost-effective analysis can be productively used in a practical, easily understood comparison of alternatives. Review of either the extensive literature or the numerous examples of systems analysis is intentionally avoided. Instead the discussion delineates an equitable way to develop cost measures, to assess achievement (as one outcome) measures and to compare instructional program alternatives using these measures.

The purposes of demonstration programs are not only to prove whether or not specific methods or practices work and to promote their implementation on a broader basis, but also to increase the understanding of the consequences of possible alternative policy decisions. To accomplish these purposes, there must be, at least, an adequate description of

the alternative, its resource requirements, its cost, and the outcomes achieved. It should be fairly obvious that additional information is needed before a final assessment can be made. For example, special circumstances that may have contributed to success should be identified--even though they may not be quantifiable; other measures of outcome may be more important than the specific measures selected for a particular one-time analysis and should be identified.

Historically, the success record of these programs, as measured by the adoption of the new ideas, is not outstanding. Demonstration programs tend to remain separate entities, observed and studied until the outside funding ceases. Their future might be more promising if the data they hold can be extracted through analysis.* It is true that, at this time, there is insufficient evidence that a better quality of information will improve the effectiveness of demonstration programs as a policy analysis tool. On the other hand, a reasonable assumption, in the absence of this evidence, is that improved information should strengthen the basis on which judgments are made in policymaking.

The usual practice in analyzing demonstration programs is to focus on determining whether or not a program is effective--that is, How well does the program accomplish its purpose? This is, of course, commendable. When coupled with an attempt to compare one program with other alternatives, the practice is even more commendable.

---

*Guidelines for analyzing demonstration programs and disseminating information are discussed in: *Increasing the Effectiveness of Educational Demonstration Programs*, Sue A. Haggart and Marjorie L. Rapp, The Rand Corporation, R-1120, August 1972.

Analysis in the comparative mode provides necessary inputs to policymaking. Thus, the emphasis in this discussion is on analysis to facilitate the comparison of alternative programs.

In the comparative analysis of alternative programs, there is a need to know more about each program and to know exactly what the numbers usually given to describe a program really mean. In short, the typical cost per student and the outcome reported in average achievement gain are not adequate. This limited information does not provide a sufficient basis for assessing the merit of alternative programs.

COST BASIS FOR COMPARING PROGRAMS

Any examination of alternative programs must be concerned with both their cost and their effectiveness. The current state-of-the-art in costing educational programs does not provide a *comparable* basis for evaluating alternative programs. The usual practice is to give the cost per student for a program with little or no indication of what is included in the cost.

When the cost per unit of achievement is used, both the cost and the effectiveness measurement problems are severe. *Education Turnkey News* has drawn attention to several aspects of using this ratio:

> Even when accurate costs are obtained, it is difficult to compare them with school costs to see which is less, since school costs are kept and reported differently. The comparisons may reveal nothing more than different figures, especially since the firms [performance contractors in the context of this quotation] may depreciate certain items much more rapidly than schools... It is even more difficult to try to contrast effectiveness with cost. If effectiveness is reported in tenths of a year's achievement, which some statisticians feel is cutting it too closely, and that figure is divided into cost data which is part hidden and part hypothetical, what does the public get? Will a school board really base a major decision on curricular changes on such a "cost per unit of achievement" figure?*

---

*Martin Reed and Peter Briggs, *Education Turnkey News*, Education Turnkey Systems, February-March 1971, p. 11.

The ratios of cost per student and of cost per unit of achievement are widely used, probably because of the false confidence these "numbers" engender and the relative ease with which they can be generated. In most instances, the cost per unit of achievement ratio masquerades as the output of cost-effectiveness analysis.

Wisely used, cost-effectiveness analysis of alternative programs produces several outputs--the aspects of cost, the measures of effectiveness, and the *relationships between cost and effectiveness*. Only very seldom is a ratio of cost per student or cost per unit achievement the appropriate end result of a cost-effectiveness analysis.

In estimating the program cost to be used in *comparing programs*, all resources needed for the program are included; that is, the resources already available within a specific district or assets inherited from discontinued programs are included in the calculation of resource use. A standard price for common resources, such as teachers, is used. The resulting estimated program cost is identified as the *comparable replication* cost. It is, in essence, an "index cost" that permits comparisons of programs on an equitable basis.*

In estimating the program cost to be used in deciding whether or not a particular program should be implemented in a specific district, the resources available within the district and district-specific prices for these resources must both be determined. The resulting estimated program cost in this case is the *incremental cost* to the district.

---

*The details of the methodology are discussed in *Program Cost Analysis in Educational Planning*, S. A. Haggart, The Rand Corporation, P-4744, December 1971.

Program cost analysis provides much of the information needed by the district in making the decision about whether or not to plan an implementation of a demonstration program. More detailed analysis helps determine the configuration that can be afforded within the resource constraints of the district. Two points should be made clear. First, these cost estimates are *planning* cost estimates. Much greater detail and accuracy are required to meet the needs of actual implementation and financial accountability. Second, analysis of the dollar cost alone does not provide adequate information for educational decisions; for this reason, the analysis of both the dollar and non-dollar resources for alternative programs is necessary.

ANALYSIS OF OUTCOME

Demonstration programs funded by federal or state agencies are usually required to administer pre- and post-tests in academic subjects so that the gain in achievement attributable to the program may be used as one measure of program success. Neither the administrators responsible for these programs, nor the evaluators are very satisfied with this crude measure of program success. Evaluators of most specially-funded programs do other types of analysis that yield more useful data for internal program use. The average gain of a sizable group of students does not tell you whom the program is reaching, where the program is in need of improvement, or where the program is having outstanding success.

A possible solution to the program is to use individual expectancy and observed scores to develop a measure for judging program success. Individual scores are aggregated to provide a measure of program success

directly related to the expected achievement of the students in the program.

The assumption is made that a student's pre-test, given before the special program starts, represents the average rate of gain made during the previous years of schooling. It is further assumed that with no special treatment the same *average* gain will be made in future years. (Caution should be exercised in using this measure in the first two or three grades because there would not have been sufficient time for random variations in learning rate to average out.) While a constant rate of gain over the years for an individual student cannot be assumed, the average rate of gain over the past years appears to be a reasonable predictor of the average gain during future years.

The difference between the program's expected gain and achieved gain at the end of the program, can be reasonably attributed to the effect of the program. A t-test on the difference between the expected and observed means can be used to determine whether the difference is significant.

Achievement data can be displayed graphically in such a way as to be useful as a management tool. An achievement idiograph can be constructed for subjects of the demonstration program.* The x-axis shows individual programs. The y-axis shows grade levels. The idiograph shows the pre-test grade level, the expected gain, and the post-test grade level. It is possible at a quick glance to ascertain which programs are making better than, or less than, expected gains.

**Idiographic Analysis of Achievement Measures*, M. L. Rapp and S. A. Haggart, The Rand Corporation, P-4880, August 1972.

Idiograph achievement analysis is useful for across-program comparisons of the relative success of a group of demonstration programs, all of which are using different techniques, materials, and classroom organizations to achieve a common goal. By automatically taking into account differences among program populations, this technique results in each program's being evaluated in terms of its own expectancy. It also obviates the need to make adjustments for differences in the meaning of test scores.* The achievement idiograph for individual reading demonstration programs is shown in Fig. 1.

Idiographic achievement analysis provides a basis for answering the question: "What kinds of results can we reasonably expect from a demonstration program?" At the same time, it removes the necessity for control groups which are statistically inappropriate when students are not randomly assigned to experimental or control classes. In short, idiographic achievement analysis provides a basis for comparing programs on an equitable basis.

THE STRUCTURE FOR COMPARING PROGRAMS

A practical approach to comparing programs involves the use of cost-effectiveness analysis. Anything more than a brief mention of the techniques of comparing cost and effectiveness, however, is

*The Educational Testing Service is scheduled to publish the results of their research effort designed to translate a student's test score on any of the seven standardized reading tests into a score on any of the other tests.

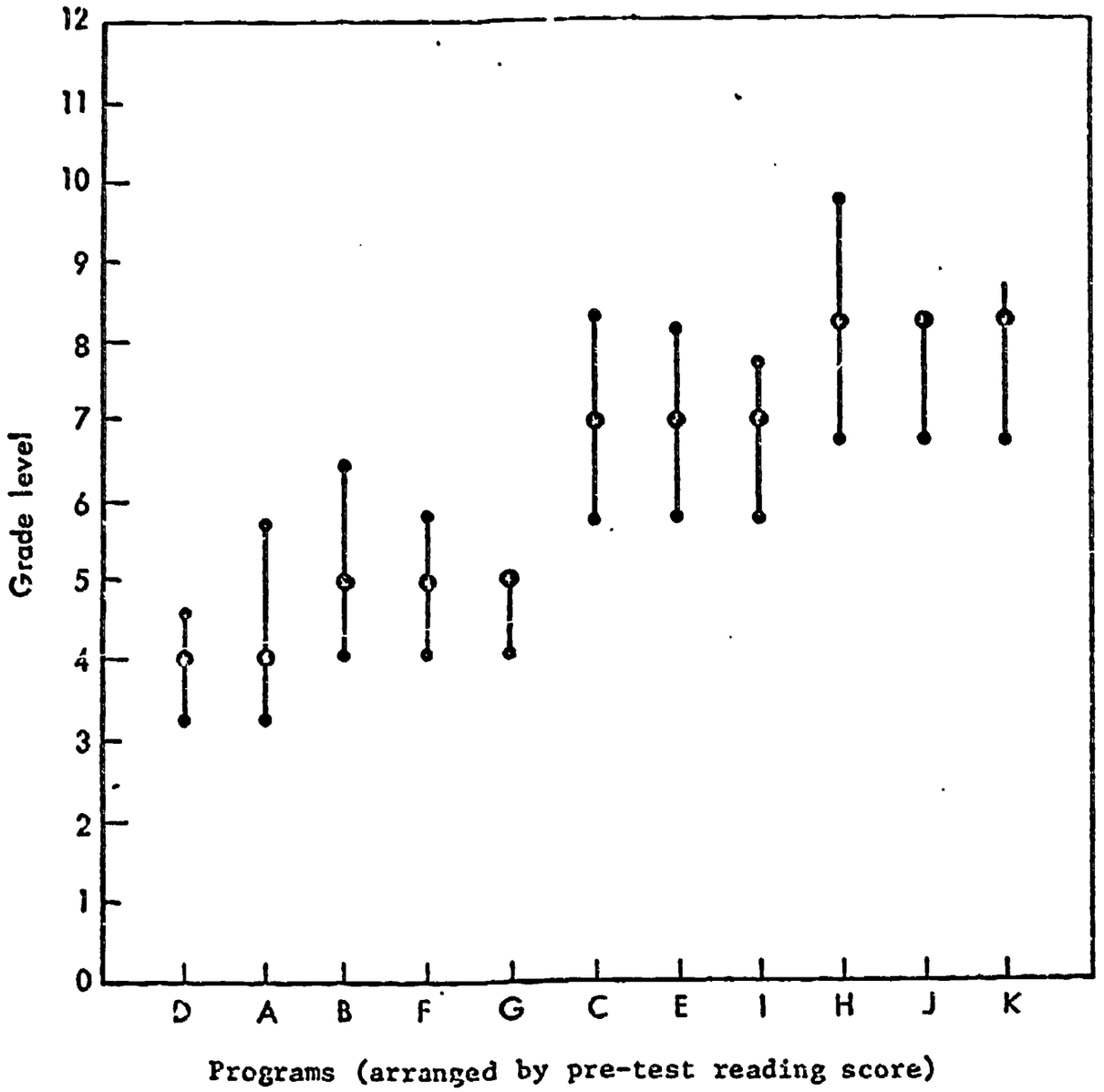

Fig. 1—Idiographic achievement analysis by program

beyond the scope and intention of this discussion.* Here the concern is mostly with the appropriateness of the cost and outcomes measures and with a fair comparison of alternative programs using these measures.

In comparing alternative programs, information is needed about the program per se. What are the objectives of the program? What target population was the program designed to reach? How many students were in the program? The answers to these questions, along with the information about program cost and outcome, provide the basic data for comparing programs.

Programs can be compared without using the analytically risky cost-effectiveness ratio *alone*. Three alternatives are available within the limitations of today's state-of-the-art in education planning.

The *equal-effectiveness approach* groups programs by their outcome and then ranks programs within each outside group by the cost to achieve that level of outcome. Conversely, the *equal-cost approach* examines the different levels of outcome achieved by each program with equal, or nearly equal, cost. That is, programs can be ranked by the outcome achieved for a given cost. A third approach is *pair-wise comparison*. In this procedure, the cost and outcome of two alternative programs are compared, and the value judgment of the decisionmaker comes into play in deciding whether or not the better outcome is worth the additional cost.

---

*There is an extensive body of literature on this subject. For some of the better discussions, the reader is referred to Gene H. Fisher, *Cost Considerations in Systems Analysis*, American Elsevier Publishing Company, New York, 1970, and Roland N. McKean, *Efficiency in Government Through Systems Analysis: With Emphasis on Water Resources Development*, John Wiley and Sons, Inc., New York, 1958.

Care must be taken, however, not to rely solely on cost-effectiveness numbers on making comparisons. This is, of course, true in any decisionmaking context, but it is especially important in educational planning. The measurement of achievement is imprecise enough so that decisions should not be based on relatively small differences in outcome. Just what degree of difference warrants a change is a matter for each decisionmaker to determine in the light of other information.

The program's comparable cost measures and the idiographic analysis of achievement described above provide a greatly improved data base for comparing the cost and outcome of alternative programs.

In the cost area, the data base is improved by developing several aspects of cost: the acquisition cost, the operational cost, the comparable replication cost. In the outcome area, the idiographic analysis of achievement permits the planner to examine the actual outcome in the light of the expected outcome for a specific group of students and to identify the performance of particular subgroups of students within the program.

Together, these measures allow the policymakers to examine the multidimensional aspects of both cost and outcome and lessen the reliance on a single number representing either *the* cost or *the* effectiveness of a program.

COMPARING THE COST AND OUTCOME

Brief examples will illustrate the use of the equal-effectiveness, equal-cost, and pair-wise comparison of alternative programs.

To review for a moment, the policymaker has available a description of the program (including the number of students), the estimates of the comparable replication cost, the outcome (achievement gain) and the idiograph of achievement data. It may be that the policymaker and the analyst are the same person or at least on the same organizational level. In reality, of course, this is not likely to be the case; there will be groups of analysts and decisionmakers, more often than not organizationally separated. The role of the analysts is to provide the results of analysis along with enough information about the structure of the analysis and the underlying assumptions and data so that the policymakers are "well-informed" and hence have a better basis for exercising their judgment.

To achieve the goal of showing the use of cost-effectiveness analysis in comparing programs, illustrative summary information about several hypothetical demonstration programs shown in Table 1 is used. For each program, an idiograph of achievement data and supporting cost analyses would be available. The resource requirements for staff, equipment, special facilities, materials, and training would be defined as part of the program description and the cost analyses.

The acquisition cost, the two-year operating cost, and the total cost for two year's operation are given for remedial reading programs scaled for 120 students. In a real situation, all costs would be estimated as comparable replication costs. The acquisition cost varies from a low of $2,000 to a high of $9,000, reflecting differences in resources required for each program. For example, one program might use a heavily equipped resource center, while another program would emphasize pre-service teacher training or student materials. The operating cost also varies from a low of $11,000 to a high of $41,000. The

Table 1

SUMMARY OF PROGRAM COST AND OUTCOME INFORMATION

Comparable Replication Cost[a] ($ thousands)				Outcome Data[b]				
Program[c]	Acquisition	Operational	Total	Pre-test Level[d]	Expected Gain	Observed Gain	Difference Gain	Post-test Level
A	2	41	43	3.2	0.8	2.6	1.8	5.8
B	6	34	40	4.0	1.0	3.0	2.0	7.0
C	3	13	16	5.7	1.3	2.7	1.4	8.4
D	2	13	15	3.2	0.8	1.5	0.7	4.7
E	4	38	42	5.7	1.3	2.5	1.2	8.2
F	2	22	24	4.0	1.0	1.8	0.8	5.8
G	3	14	17	4.0	1.0	1.0	---	5.0
H	6	16	22	6.6	1.6	3.2	1.6	9.8
I	2	12	14	5.7	1.3	1.9	0.6	7.6
J	2	23	25	6.6	1.6	1.6	---	8.2
K	9	35	44	6.6	1.6	2.0	0.4	8.6

[a]Actual programs were for varying numbers of students. The comparable replication cost was estimated for the *same number of students*.

[b]Achievement gain over a two-year program. All outcomes expressed as mean grade equivalents.

[c]Remedial reading programs; all costs for 120 students in each program.

[d]Entry grade level in eighth grade.

tradeoff of lower acquisition cost versus higher operating cost might be a deciding factor in a district's capability to implement a program.*

The important characteristic of these cost measures is their *comparability*. The decisionmaker is fully aware not only of *what is included* in each estimate, but also of the fact that *standard prices* were used for the resources required. Moreover, because these illustrative programs are scaled to be alternatives, the cost-effectiveness ratios have some meaning in the initial ranking of the programs.

In assessing the outcome, the decisionmaker has knowledge about the target population reached by the program as well as the mean achievement gain. The differential gain is the observed gain, less the *expected gain* for a specific target population.

Use of the cost and outcome information of Table 1 results in the ranking of programs shown in Table 2. Programs C, H, I, and D, with lower costs per unit of achievement, rank in the top four, while Programs A, E, G, and K have the highest cost per unit of achievement. The composition of "the better" four programs changes somewhat if outcome (either expected or differential) is the basis of ranking.

As shown in Table 3, consideration of gain alone puts Programs H, B, A, and C in the top four; Programs A and B are added to the top grouping, and Programs I and D rank considerably lower (seventh and tenth for observed gain, and seventh and ninth for differential gain). Programs G and K seem to be consistently low in the rankings.

---

*See Sumner's discussion in R-955/2, *A Guide to Education Performance Contracting--Technical Appendix*, Sue A. Haggart, G. C. Sumner, and J. Richard Harsh, The Rand Corporation, March 1972.

Table 2

RANKING OF PROGRAMS BY THEIR COST-EFFECTIVENESS RATIOS

Program	Comparable Replication Cost ($ thousands)	Observed Gain (years)	Cost/ Effectiveness Ratio	Program Ranking
A	43	2.6	16.5	8
B	40	3.0	13.3	5
C	16	2.7	5.9	1
D	15	1.5	10.0	4
E	42	2.5	16.8	9
F	24	1.8	13.4	6
G	17	1.0	17.0	10
H	22	3.2	6.9	2
I	14	1.9	7.4	3
J	25	1.6	15.6	7
K	44	2.0	22.0	11

Table 3

COMPARISON OF PROGRAMS RANKED ON OUTCOME ONLY

Ranked by Observed Gain	Ranked by Cost-effectiveness Using Observed Gain	Ranked by Differential Gain
1 H 3.2	C	B 2.0
2 B 3.0	H	H 1.6
3 A 2.6	I	A 1.6
4 C 2.7	D	C 1.4
5 E 2.8	B	E 1.2
6 K 2.0	F	F 0.8
7 I 1.9	J	D 0.7
8 F 1.8	A	I 0.6
9 J 1.6	E	K 0.4
10 D 1.5	G	G —
11 G 1.0	K	J —

What happens if we *group* programs *by nearly equivalent outcomes* and select the lowest cost program within that group? Conversely, what programs are identified as the most effective within a group of programs whose *cost is equal or nearly equal?* These results are

shown in Tables 4 and 5. In the equal-effectiveness case, Program H is more cost-effective than Program B in achieving a mean observed gain of approximately 3.0 years.

When programs are ranked by outcome within nearly equal cost groups (Table 5), Programs B, H, and C are the most cost-effective for a given cost. Here again, Program K really comes off poorly: the program is more costly than any other program, ranks fairly near the average outcome of all Programs A through K, and produces results on a par with four of the seven lower cost programs.

A useful and revealing technique is *pair-wise comparison* of two alternatives. It is especially revealing in that a wider view of the problem and the alternatives is promoted.

Briefly, in a pair-wise comparison, the outcome and cost are examined in relation to each other, and the value judgment of the decisionmaker is exercised in deciding if the additional outcome is worth the additional cost.

In the example, Programs B and F address the same target population (the same average pre-test entry level). Program B has a cost of $40,000 and an achievement gain of 3.0. Program F, on the other hand, with a gain of 1.8, has a cost of only $24,000:

	Cost	Outcome	Cost/ Effectiveness
Program B	$40,000	3.0	13.3
F	24,000	1.8	13.4

Table 4

PROGRAM RANKING BY COST WITHIN GROUPS OF NEARLY EQUAL EFFECTIVENESS

Observed Gain					
3.2 -- 2.5			2.0 -- 1.0		
Program C	$16,000	(2.7)	Program I	$14,000	(1.9)
H	22,000	(3.2)	D	15,000	(1.5)
B	40,000	(3.0)	G	17,000	(1.0)
E	42,000	(2.5)	F	24,000	(1.8)
A	43,000	(2.6)	J	25,000	(1.6)
			K	44,000	(2.0)

Table 5

PROGRAM RANKING BY OUTCOME WITHIN GROUPS OF NEARLY EQUAL COST

Cost Groups								
$44,000-$40,000			$25,000-$20,000			$19,000-$12,000		
Program B	3.0		Program H	3.2		Program C	2.7	
A	2.6		F	1.8		I	1.9	
E	2.5		J	1.6		D	1.5	
K	2.0					G	1.0	

The decisionmaker must ask if the additional 1.2 achievement gain is worth $16,000.

The pair-wise comparison is particularly enlightening in view of roughly the *same cost-to-effectiveness ratio* for both programs. This illustrates how the cost-effectiveness ratio, by itself, can be misleading; an indifference on the part of the decisionmaker is implied--an indifference not at all evident from the dimensions of cost and outcome.

These brief examples illustrate a simple, analytically sound way to make the most of cost and outcome data in comparing alternative programs. It is possible to effectively rank or select programs matched to both educational needs and available resources. Because the multi-dimensional aspects of cost and effectiveness are known and can be arrayed in a straightforward fashion, reliance on the single, enigmatic cost per student or cost per unit of achievement can be avoided.

Because the analytical basis of comparing programs is similarly straightforward, the results are likely to be more credible to the policymaker. The overall impact should be an enhancement of policy-making due primarily to an increase in the quality of information on which judgments are made and a decrease in the gap between the policymakers and the practitioners who implement the policy decisions.